

International Life Skills Survey (ILSS)
All-Hands Meeting
Washington, D.C.
January 20-23, 1999

Summary Notes

Participants

Ronald Anderson	Graham Lowe
David Baker	Julie McAuley
Kirk Basehore	Barry McGaw
Marilyn Binkley	Scott Murray
Anna Borkowsky	Lynn Offermann
Axel Buchner	Eugene Owen
Michael Campion	Jean Pignal
Judith Ebach	Jean-Paul Reeffer
Iddo Gal	Hermann Wolf Richter
Elena Grigorenko	Dominique Simone Rychen
Mieke van Groenestijn	Eduardo Salas
Raymond Hanson	Laura Salganik
Lisa Horvath	Mary Jane Schmitt
Christina Kary	Eftychia Sidiropoulou
David Kastberg	Robert Sternberg
Irwin Kirsch	Dave Tout
Eckhard Klieme	Albert Tuijnman
John Konstant	Trevor Williams
Qiwu Liu	Kentaro Yamamoto

Introduction and Context for the Study

Marilyn Binkley welcomed everyone and gave a brief overview of the goals of the meeting. Albert Tuijnman detailed the history of IALS and ILSS along with the OECD structure and explained the position of ILSS within that structure and within the INES Program. The ILSS organizational chart was discussed, and it was noted that representatives from the development teams to the Project Coordination Committee are needed. Discussion followed on the relationship of ILSS to PISA, the labor-market orientation of ILSS, and the constituencies that ILSS will target.

Scott Murray gave a presentation on the policy interests surrounding ILSS. The current design parameters of the assessment were outlined along with the quality assurance and dissemination strategies. It was noted that producing individual-level scores is beyond the current ILSS design capacity. The option for specific case studies on a per-country basis was discussed.

Marilyn Binkley presented the current notions of an overarching conceptualization of life skills, and a discussion followed on the position of the current domains within the larger framework. Participants relayed specific outcomes that they hope ILSS will result in (See Annex A).

The Frameworks

A discussion of the frameworks was initiated through participant questions regarding the individual frameworks (See Annex B). Teams were encouraged to take these concerns into account in their deliberations and framework revisions. It was noted that the feasibility studies are meant to provide evidence for the conceptualizations outlined in the frameworks. Each development team gave a short presentation on the conceptualization of their domain and updated the group as to their current work.

Discussions followed each presentation, touching on a variety of issues. The unique nature of the evolution of the literacy framework as it developed over the last decade was discussed along with possible lessons that can be applied from it to the other frameworks. The notion of correlation to general intelligence (*g*) was discussed in regard to several of the skill domains along with the scaling and scoring rubrics to be used for each domain. Issues surrounding time constraints and shortening the assessments were discussed, as were possible statistical linkages to PISA for the literacy and numeracy domains. It was noted that the circulated problem-solving materials are working drafts, and are not intended for distribution. Utilizing the feasibility studies to examine relationships across scales was discussed as well.

The purpose of the revised frameworks was discussed, along with the various options for organizing the resulting documents, such as in a publication or as a revised briefing book. The implications associated with finalizing and publishing the frameworks were detailed as well. It was decided that the revised frameworks will be included in a revised briefing book/prospectus. A paper by Irwin Kirsch on framework structure and construction will be circulated as a reference for revisions.

Presentations

Robert Sternberg and Elena Grigorenko presented the results of their feasibility study conducted in the U.S. and in Spain. A general discussion of the results and of the practical cognition scale followed, touching on such issues as scaling, scoring, and item difficulty.

Judith Ebach, Eckhard Klieme, and Eftychia Sidiropoulou presented the results of their feasibility study conducted in Germany. A general discussion of the results and of the problem-solving scale followed, including the nature of problem solving, the nature of the tasks involved in the assessment, the length of the assessment, the scoring typology, and correlations with other domains.

Measurement Issues

Irwin Kirsch led a discussion on measurement and design issues. The three primary design criteria—relevance, comparability, and interpretability—were given for an assessment such as ILSS. Translation, transportability, and comparability were discussed, and set procedures will be developed and disseminated. In terms of scaling, it was noted that while IRT is the presumed method at this point, it might not be ideal for all of the domains. Teams must decide how scores will be created for their domains for reporting purposes and what construct areas will be used for interpretation. For these reasons, teams must have a complete understanding of what is being measured and what “positioning along the scales” means. When considering these issues, teams must consider political criteria along with the scientific. At the present time, discriminate validity will not be available until the main assessment in 2002, and not across all of the scales.

In terms of interdependency between the scales, there is a possibility that modeling and examining items across the scales could be done. But as this is not currently possible on such a large scale, a significant amount of time and work would need to be invested to successfully implement such a design.

Increased coordination and communication among the development teams were discussed as well, based on conceptual overlap among the domains. The need for further communication regarding this was noted as was the need for a clear overarching conceptualization of life skills.

Timeline and Feasibility Studies

A timeline was introduced whereby feasibility studies would be conducted in Spring/Fall 1999, the pilot assessment in Spring 2001, and the main assessment in Spring 2002. The Project Management is working in conjunction with OECD on securing the commitment of 10 countries for the main assessment.

The feasibility studies would consist of 200-400 cases per country and would need to be done in at least two languages. The Project Management Team will work with the development teams on logistics regarding the studies, and any comments or suggestions regarding the feasibility studies should be relayed to the Project Management Team. The development teams must decide what empirical data they require and design the feasibility studies accordingly.

Critical Item Elements

Marilyn Binkley led a discussion on cultural transportability and the application of the variables relating to task difficulty in the literacy framework. The need to accurately define critical elements related to each item was discussed, in particular as it relates to securing accurate item translation. In addition, participants were advised to be mindful of

various potential hazards relating to translation including task-load changes, redundancy level, embedded text, punctuation, graphics, and paper sizing.

Wrap-Up

Over the course of the meeting the development teams met privately for several working sessions. When the teams reconvened together for the final group session each team reported on the progress made at the All-Hands meeting and the next steps the teams will take.

Participants were advised to include rationalizations along with any item requests for the background questionnaire as there is very limited space available and the developers will need to prioritize. However, development teams were encouraged to provide as much input as possible to the background questionnaire development team to inform their deliberations.

Development teams are encouraged to increase communication through the use of such means as listservs, which can be set up easily upon request. A contact person from each development team must be named for management purposes. This contact person will participate in a teleconference call in the first week of February with the Project Management Team (See Annex C for the results of the call). Decisions must be shared with the Project Management Team and archived. The Project Management Team will prepare guidelines regarding communication procedures and will provide further clarification regarding the feasibility study reporting process and framework revisions.

Revised frameworks should be delivered by March 15 if possible. In addition, a paragraph regarding the policy relevance of the skill domains and a brief description (1 page) of achievements and decisions coming out of the All-Hands meeting should be delivered as well.

Participants relayed their short-term goals for ILSS to the group, and it was noted that ILSS is now moving into an implementation phase wherein coordination and communication among participants is key to the success of ILSS.

Comments, suggestions, and materials for the Project Management Team should be sent to John Konstant (email: jkonstant@air-dc.org).

Annex A: Potential Outcomes as Identified by Participants

- Horse race and sub races
- How different skills...
- Skills by education
- Organized link to scientific community
- Projected interactive world of the future
- Relocation information
- Business/industry information
- In-school/out-of-school information
- Value added information, justification through new data
- New framework updating conventional wisdom
- How the domains relate to one another
- Returns to these skills in labor market, (i.e., socially, quality of life, and general health (mental) questionnaire)
- feedback into theory development for the domains
- Sub-populations--equity?
- Feedback into training and development
- Coping with novelty
- Increasing measurements in creativity
- Correlation with health-related factors (e.g., child mortality, marriage)
- The measures/confidence in results, legitimacy
- Data quality frame
- Pre-field data report to let NPMs know what is wanted
- Literacy, Numeracy—practices and impact on maintaining and strengthening skills, and home practices and their impact
- Variation in how skills relate to education and background
- Broadening the domain to include new skills
- Better understanding of the contextual barriers and facilitators to the use and development of these skills
- Links between skills and job satisfaction
- Ability of disabled population
- Unemployed/low education groups information, practical implications of the study
- Cross-cultural perspectives
- Outline of decision-making model for policymakers
- Guidelines on how to use this data and how to train individuals in life skills, based on the theoretical and empirical evidence
- Link with intangible assets community
- Clear notion of avenues of access or lack of access to skills
- Relative independent contribution of each skill to outcomes
- Interpretability of the measures: what are we measuring/teaching as outcome of the levels?
- Street versus school learning behavior
- Formal versus informal reasoning strategies

- Retention skills over ages
- Continued re-education process of macroeconomists
- Expand network of like-minded people
- LAST MEETING IN THE SUN!

Annex B: Participant Questions and Comments Regarding the Frameworks

Literacy:

- What will be new?
- Don't need the old understanding as much, relevance disappearing
- Overlap with Numeracy

Numeracy:

- Would you be able to apply the Numeracy scale to Problem Solving and obtain more information?
- How will the data be reported?
- Overlap with Literacy
- Overlap with Problem Solving
- Attitudes, how do they fit in, and how do you deal with the link to performance?

Problem Solving:

- In the end, how do you differentiate from Practical Cognition?
- Would you be able to apply the Numeracy scale to Problem Solving and obtain more information?
- Isn't context a problem?
- Non-formal problem solving not covered
- Overlap with Numeracy
- Contextual overlap with Teamwork
- The cognitive demands on respondents appear very high
- Potential problem with non-respondents for these questions
- Generalizability is a question
- What about algorithmic thinking? (e.g., flowcharts and such)

Practical Cognition:

- How will the scaling be done, the norms set, what experts are involved?
- What are the theoretical underpinnings, and how does it relate to other models like expert knowledge and such?
- In the end, how do you differentiate from Problem Solving?
- Appears to be normative and not task specific
- Culture specific knowledge on scaling and scoring
- Why are only work settings utilized?
- Culturally specific in terms of item development
- Context bound

Teamwork:

- How would you score the scale?
- How do you deal with communication skills?
- Workplaces are used, but what about other team settings like community participation, collaborative learning, and non-formal settings?

- How do you deal with social desirability?
- Attitude survey, inferences to behavior, need explicit linkages to behavior somehow
- Culture specific knowledge on scaling and scoring
- Culturally specific in terms of item development

Information and Communication Technology:

- Attitudes, how do they fit in, and how do you deal with the link to performance?
- How do you bring this scale back to other scales?
- What is the definition itself, and how do you justify it?
- How do you plan to cope with a fast changing world?
- How do you integrate across borders?
- If it is literacy, then it needs a theoretical rationale for why it is literacy, and how it is used in relation to other literacies.

All:

- Training interviewers
- Discriminate validation
- The models are multi-dimensional, but the scales are uni-dimensional, what does that mean for scaling?
- How do you account for think-aloud protocols, tape recording and such?
- How do you plan to cope with a fast changing world?
- How the domains are related to each other, reflect and discuss overlap explicitly
- Cultural bias in terms of defining Like Skills in this manner
- Ethics and values questions, privacy questions
- Unintended constructs, *g*-loaded, situational judgement and context specific
- Can the constructs be separated from culture, particularly Teamwork and Practical Cognition?
- If not specific enough, aren't Problem Solving and Practical Cognition being forced away from what they are trying to measure?
- General conception of the field, need to justify it conceptually, and where overlap exists
- In terms of reporting levels for the different skill domains, particularly ICT and Teamwork, where does difficulty come in?
- Will things be metric and comparable enough to be reported? It has to be decided if it will just be done as behavioral reports, don't want to be arbitrary
- How do you integrate across borders, especially in terms of ICT?

Annex C: Results of the February 3, 1999 Conference Call

A conference call was held on February 3, 1999 with members of the Project Management Team and several of the development team contact persons.

Scott Murray briefed the participants on the meeting held at ETS the prior week. It was decided that items from one scale will not be used to “borrow information” for other scales. Analyses will be done on the pilot data to see what kind of relationships exist among the scales, but for design purposes, development teams should assume that they will not obtain any information on their domain from the other scales. The Project Management Team also noted a clear overlap among the Literacy, Numeracy, and Problem Solving frameworks. While there is no solution to this problem right now, the Project Management Team will continue to work with the development teams to resolve it. Finally, a provisional schedule was outlined wherein the feasibility studies would conclude by July 15, 1999. The analyses would be finished by August 31, 1999 and a design meeting would be held in September.

Over the course of the conference call it was decided that each development team would have 3 liaisons to the Project Management Team to deal with project management, including contracts, scale development, and technical aspects of the survey, in conjunction with ETS. For Numeracy, Irwin Kirsch will serve as the development and technical liaison and Julie McAuley as the management liaison. Irwin Kirsch will also serve in the same role for the Problem Solving Team along with Marilyn Binkley who will be the management liaison. For Teamwork, Marilyn Binkley will serve as the management and development liaison, and Don Rock as the technical liaison.

The participants discussed increased communication between Irwin Kirsch and the Numeracy and Problem Solving teams regarding the overlap issue. In addition, teams are encouraged to consider the use of incentives and how they think it would impact their scale.

The March 15, 1999 deadline for frameworks was re-affirmed as well.